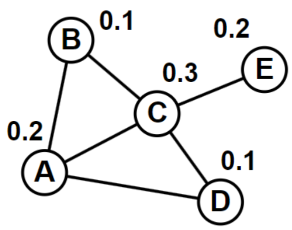
Final Exam (Graph Neural Networks –Fall 2023)

Full Name:
Student ID:

1. Consider an undirected graph G of five nodes A, B, C, D, and E given in the following figure. Each node has initial features that are the numbers standing next to it (i.e., the initial feature of node ‘A’ is $h_A^{(0)} = 0.2$). According to GraphSAGE model with a AGGREGATE is a MEAN function, the feature of a node i at layer k can be updated as:



$$h_{N(i)}^{(k)} = \text{AGGREGATE}\left(\left\{h_u^{(k-1)}, \forall u \in N(i)\right\}\right)$$

$$h_i^{(k)} = \text{ReLU}\left(h_i^{(k-1)} \,\|\, h_{N(i)}^{(k)}\right)$$

where $\|$ is a concatenation, $\text{ReLU}(x) = \max(0, x)$, $N(i)$ is the neighbour nodes of node i.

   a) Calculate the feature of each node at $k = 1$.
   b) Calculate a graph-level embedding h_G by using a ‘Mean’ global pooling when $k = 1$.

SOLUTIONS:

$$h_{N(i)}^{(k)} = \text{AGGREGATE}\left(\left\{h_u^{(k-1)}, \forall u \in N(i)\right\}\right)$$

$$h_{N(A)}^{(1)} = MEAN\left(h_B, h_C, h_D\right) = MEAN\left(0.1, 0.3, 0.1\right) = 0.17$$

$$h_{N(B)}^{(1)} = MEAN\left(h_A, h_C\right) = MEAN\left(0.2, 0.3\right) = 0.25$$

$$h_{N(C)}^{(1)} = MEAN\left(h_A, h_B, h_E, h_D\right) = MEAN\left(0.2, 0.1, 0.2, 0.1\right) = 0.15$$

$$h_{N(D)}^{(1)} = MEAN\left(h_A, h_C\right) = MEAN\left(0.2, 0.3\right) = 0.25$$

$$h_{N(E)}^{(1)} = MEAN\left(h_C\right) = MEAN\left(0.3\right) = 0.3$$

$$h_i^{(k)} = \text{ReLU}\left(h_i^{(k-1)} \,\|\, h_{N(i)}^{(k)}\right)$$

$$h_A^{(1)} = \text{ReLU}\left(h_A^{(0)} \,\|\, h_{N(A)}^{(0)}\right) = Max\left(0, [0.2, 0.17]\right) = [0.2, 0.17]$$

$$h_B^{(1)} = \text{ReLU}\left(h_B^{(0)} \,\|\, h_{N(B)}^{(0)}\right) = Max\left(0, [0.1, 0.25]\right) = [0.1, 0.25]$$

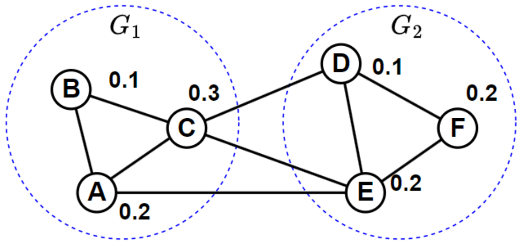
$$h_C^{(1)} = \text{ReLU}\left(h_C^{(0)} \,\|\, h_{N(C)}^{(0)}\right) = Max\left(0, [0.3, 0.15]\right) = [0.3, 0.15]$$

$$h_D^{(1)} = \text{ReLU}\left(h_D^{(0)} \,\|\, h_{N(D)}^{(0)}\right) = Max\left(0, [0.1, 0.25]\right) = [0.1, 0.25]$$

$$h_E^{(1)} = \text{ReLU}\left(h_E^{(0)} \,\|\, h_{N(E)}^{(0)}\right) = Max\left(0, [0.2, 0.3]\right) = [0.2, 0.3]$$

$$h_G = MEAN\left(h_A^{(1)}, h_B^{(1)}, h_C^{(1)}, h_D^{(1)}, h_E^{(1)}\right) = [0.18, 0.224]$$

2. Consider an undirected graph G of six nodes A, B, C, D, E and F given in the following figure. The graph G contains two cluster G_1 and G_2. Each node has initial features that are the numbers standing next to it. According to ClusterGCN model, the feature of a node i at layer k can be updated as:



$$h_{N(i)}^{(k)} = \text{MEAN}\left(\left\{h_u^{(k-1)}, \forall u \in N(i), G_u = G_i\right\}\right)$$
$$h_i^{(k)} = \text{ReLU}\left(h_i^{(k-1)} \,\|\, h_{N(i)}^{(k)}\right)$$

where $\|$ is a concatenation.

Calculate the output representations of all nodes at layer $k = 1$.

SOLUTIONS:

$$h_{N(i)}^{(k)} = \text{AGGREGATE}\left(\left\{h_u^{(k-1)}, \forall u \in N(i)\right\}, G_u = G_i\right)$$
$$h_{N(A)}^{(1)} = MEAN\left(h_B, h_C\right) = MEAN\left(0.1, 0.3\right) = 0.2$$
$$h_{N(B)}^{(1)} = MEAN\left(h_A, h_C\right) = MEAN\left(0.2, 0.3\right) = 0.25$$
$$h_{N(C)}^{(1)} = MEAN\left(h_A, h_B\right) = MEAN\left(0.2, 0.1\right) = 0.15$$
$$h_{N(D)}^{(1)} = MEAN\left(h_E, h_F\right) = MEAN\left(0.2, 0.2\right) = 0.2$$
$$h_{N(E)}^{(1)} = MEAN\left(h_D, h_F\right) = MEAN\left(0.1, 0.2\right) = 0.15$$
$$h_{N(F)}^{(1)} = MEAN\left(h_D, h_E\right) = MEAN\left(0.1, 0.2\right) = 0.15$$

\-

$$h_i^{(k)} = \text{ReLU}\left(h_i^{(k-1)} \,\|\, h_{N(i)}^{(k)}\right)$$
$$h_A^{(1)} = \text{ReLU}\left(h_A^{(0)} \,\|\, h_{N(A)}^{(0)}\right) = Max\left(0, [0.2, 0.2]\right) = [0.2, 0.2]$$
$$h_B^{(1)} = \text{ReLU}\left(h_B^{(0)} \,\|\, h_{N(B)}^{(0)}\right) = Max\left(0, [0.1, 0.25]\right) = [0.1, 0.25]$$
$$h_C^{(1)} = \text{ReLU}\left(h_C^{(0)} \,\|\, h_{N(C)}^{(0)}\right) = Max\left(0, [0.3, 0.15]\right) = [0.3, 0.15]$$
$$h_D^{(1)} = \text{ReLU}\left(h_D^{(0)} \,\|\, h_{N(D)}^{(0)}\right) = Max\left(0, [0.1, 0.2]\right) = [0.1, 0.2]$$
$$h_E^{(1)} = \text{ReLU}\left(h_E^{(0)} \,\|\, h_{N(E)}^{(0)}\right) = Max\left(0, [0.2, 0.15]\right) = [0.2, 0.15]$$
$$h_F^{(1)} = \text{ReLU}\left(h_E^{(0)} \,\|\, h_{N(F)}^{(0)}\right) = Max\left(0, [0.2, 0.15]\right) = [0.2, 0.15]$$

3. Given a graph with an adjacency matrix A and initial node feature matrix $H^{(0)}$ as follows:

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \qquad H^{(0)} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}$$

Assume that the output of an GCNII model of all nodes at layer (k) can be calculated as:

$$H^{(k)} = \sigma\left[\left((1-\beta) I_n\right) \cdot \left((1-\alpha)\tilde{A} \cdot H^{(k-1)} + \alpha H^{(0)}\right)\right]$$

where $H^{(k)}$ denotes the output at layer k, $\tilde{A}$ is the normalized matrix ($\tilde{A}=D^{-1}A$), I_n is the identity matrix, $\alpha=\beta=0.5$, σ is a ReLU function $\mathrm{ReLU}(x)=\max(0,x)$.

a) Calculate $\tilde{A}$.

b) Calculate the output representations at layer $k=1$.

SOLUTIONS

$$\tilde{A}=D^{-1}A=\begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \end{bmatrix}$$

$$H^{(k)}=\sigma\left[\left((1-\beta)I_n\right)\left((1-\alpha)\tilde{A}\cdot H^{(k-1)}+\alpha H^{(0)}\right)\right]$$

$$\tilde{A}\cdot H^{(0)}=\begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \end{bmatrix}\cdot\begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}=\begin{bmatrix} 0.5 \\ 1.5 \\ 0.5 \\ 1.5 \end{bmatrix}$$

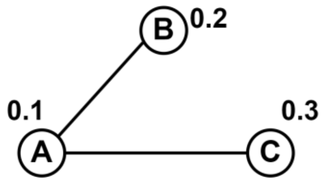
$$I_n=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\left[\left((1-\alpha)\tilde{A}\cdot H^{(k-1)}+\alpha H^{(0)}\right)\cdot\left((1-\beta)I_n\right)\right]$$

$$=\left(0.5\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\right)\cdot\left(0.5\begin{bmatrix} 0.5 \\ 1.5 \\ 0.5 \\ 1.5 \end{bmatrix}+0.5\begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}\right)=0.5\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}0.5\begin{bmatrix} 1.5 \\ 1.5 \\ 2.5 \\ 2.5 \end{bmatrix}=0.25\begin{bmatrix} 1.5 \\ 1.5 \\ 2.5 \\ 2.5 \end{bmatrix}=\begin{bmatrix} 0.375 \\ 0.375 \\ 0.625 \\ 0.625 \end{bmatrix}$$

$$H^{(k)}=\sigma\left(\begin{bmatrix} 0.375 \\ 0.375 \\ 0.625 \\ 0.625 \end{bmatrix}\right)=\begin{bmatrix} 0.375 \\ 0.375 \\ 0.625 \\ 0.625 \end{bmatrix}$$

4. Consider an undirected graph G of three nodes A, B, and C given in the following figure. Each node has initial features that are the numbers standing next to it (i.e., the initial feature of node 'A' is $h_A^{(0)}=0.1$). According to GAT model, the weight matrix W is randomly initialized as [0.5]. The feature of node 'i' at layer (k) can be updated as:



$$h_i^{(k)} = \sigma\left(\sum_{m \in N(i)} \alpha_{im} W h_m\right)$$

where: $\alpha_{im} = \frac{e_{im}}{\sum_{k \in N(i)} e_{ik}}$, and $e_{im} = \sigma\left(\mathrm{MEAN}\left(Wh_i, Wh_m\right)\right)$

σ is a ReLU function $\mathrm{ReLU}(x) = \max(0, x)$.

a) Calculate the attention coefficients e_{AB} and e_{AC}

b) Calculate the feature of node 'A' at $k = 1$.

SOLUTIONS:

$$h_i^{(k)} = \sigma\left(\sum_{m \in N(i)} \alpha_{im} W h_m\right)$$

where: $\alpha_{im} = \frac{e_{im}}{\sum_{k \in N(i)} e_{ik}}$, and $e_{im} = \sigma\left(\mathrm{MEAN}\left(Wh_i, Wh_m\right)\right)$

$$e_{AC} = \sigma\left(\mathrm{MEAN}\left(Wh_A, Wh_C\right)\right) = \sigma\left(\mathrm{MEAN}\left(0.5*0.1, 0.5*0.3\right)\right)$$

$$= \sigma\left(\mathrm{MEAN}\left(0.05, 0.15\right)\right) = 0.1$$

$$e_{AB} = \sigma\left(\mathrm{MEAN}\left(Wh_A, Wh_B\right)\right) = \sigma\left(\mathrm{MEAN}\left(0.5*0.1, 0.5*0.2\right)\right)$$

$$= \sigma\left(\mathrm{MEAN}\left(0.05, 0.1\right)\right) = 0.075$$

−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−

$$\alpha_{AB} = \frac{e_{AB}}{e_{AB} + e_{AC}} = \frac{0.075}{0.075 + 0.1} = 0.43$$

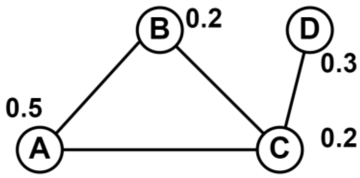
$$\alpha_{AC} = 0.57$$

−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−

$$h_A^{(1)} = \sigma\left(\sum_{m \in N(i)} \alpha_{Am} W h_m\right) = \sigma\left(\alpha_{AB} W h_B + \alpha_{AC} W h_C\right)$$

$$= \sigma\left(0.43*0.5*0.2 + 0.57*0.5*0.3\right) = 0.13$$

5. Consider an undirected graph G of four nodes A, B, C, and D given in the following figure. Each node has initial features that are the numbers standing next to it (i.e., the initial feature of node 'A' is $h_A^{(0)} = 0.5$). According to GIN model, the parameter is a fixed scalar $\varepsilon = 0.5$, the feature of a node i at layer k can be updated as:



$$h_i^{(k)} = (1+\varepsilon)\cdot h_i^{(k-1)} + \sum_{j\in N(i)} h_j^{(k-1)}$$

a) Calculate the feature of each node at $k = 1$.
b) Calculate a graph-level embedding h_G by using a 'Max' global pooling when $k=1$.

SOLUTIONS:

$$h_i^{(k)} = (1+\varepsilon)\cdot h_i^{(k-1)} + \sum_{j\in N(i)} h_j^{(k-1)}$$

$$h_A^{(1)} = (1+\varepsilon)\cdot h_A^{(0)} + \left(h_B^{(0)} + h_C^{(0)}\right) = (1+0.5)*0.5 + (0.2+0.2) = 1.15$$

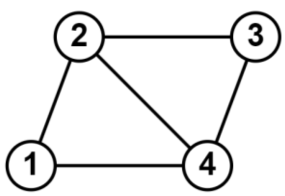
$$h_B^{(1)} = (1+\varepsilon)\cdot h_B^{(0)} + \left(h_A^{(0)} + h_C^{(0)}\right) = (1+0.5)*0.2 + (0.5+0.2) = 1$$

$$h_C^{(1)} = (1+\varepsilon)\cdot h_C^{(0)} + \left(h_A^{(0)} + h_B^{(0)} + h_D^{(0)}\right) = (1+0.5)*0.2 + (0.5+0.2+0.3) = 1.3$$

$$h_D^{(1)} = (1+\varepsilon)\cdot h_D^{(0)} + \left(h_C^{(0)}\right) = (1+0.5)*0.3 + (0.2) = 0.65$$

$$h_G^{1} = MAX\left(h_A^{(1)}, h_B^{(1)}, h_C^{(1)}, h_D^{(1)}\right) = 1.15$$

6. Consider an undirected graph G of four nodes given in the following figure. The Random Walk Positional Encoding, which is used in SAT model, of a node i can be calculated as:



$$p_i^{RWPE} = \left[\tilde{A}_{ii}, \tilde{A}_{ii}^2, \ldots, \tilde{A}_{ii}^k\right]$$

where $\tilde{A}$ is the normalized adjacency matrix $\tilde{A} = D^{-1}A$, $\tilde{A}^k$ is the k-step transition probability matrix $\tilde{A}^k = \underbrace{\tilde{A}\cdot\tilde{A}\cdots\tilde{A}}_{k}$. Calculate the positional encoding of each node at $k=2$.

SOLUTIONS

$$\tilde{A} = D^{-1}A = \begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.33 & 0 & 0.33 & 0.33 \\ 0 & 0.5 & 0 & 0.5 \\ 0.33 & 0.33 & 0.33 & 0 \end{bmatrix}$$

$$\tilde{A}^2 = \tilde{A}\tilde{A} = \begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.33 & 0 & 0.33 & 0.33 \\ 0 & 0.5 & 0 & 0.5 \\ 0.33 & 0.33 & 0.33 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.33 & 0 & 0.33 & 0.33 \\ 0 & 0.5 & 0 & 0.5 \\ 0.33 & 0.33 & 0.33 & 0 \end{bmatrix} = \begin{bmatrix} [0.33,\ 0.16,\ 0.33,\ 0.16], \\ [0.11,\ 0.44,\ 0.11,\ 0.33], \\ [0.33,\ 0.16,\ 0.33,\ 0.16], \\ [0.11,\ 0.33,\ 0.11,\ 0.44] \end{bmatrix}$$

$$p_A^{RWPE} = \left[\tilde{A}_{11}, \tilde{A}_{11}^2 \right] = [0, 0.33]$$

$$p_B^{RWPE} = \left[\tilde{A}_{22}, \tilde{A}_{22}^2 \right] = [0, 0.44]$$

$$p_C^{RWPE} = \left[\tilde{A}_{33}, \tilde{A}_{33}^2 \right] = [0, 0.33]$$

$$p_D^{RWPE} = \left[\tilde{A}_{44}, \tilde{A}_{44}^2 \right] = [0, 0.44]$$